1993

The Neutral response on attitudinal measures : an attribute of the item

Pauline Velez
San Jose State University

Order Number 1353068

The neutral response on attitudinal measures: An attribute of the item

Velez, Pauline, M.S.

San Jose State University, 1993

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

The Neutral Response on Attitudinal Measures:

An Attribute of the Item

A Thesis

Presented to

the Faculty of the Department of Psychology

San Jose State University

In Partial Fulfillment

of the Requirements for the Degree

Master of Science

by

Pauline Velez

May, 1993

APPROVED FOR THE DEPARTMENT OF PSYCHOLOGY

Howard Tokunaga, Ph.D.

Lori Bartels, Ph.D.

Steven D. Ashworth, Ph.D.

APPROVED FOR THE UNIVERSITY

M. Lou Lewandowski

ABSTRACT
THE NEUTRAL RESPONSE ON ATTITUDINAL MEASURES:
AN ATTRIBUTE OF THE ITEM

by Pauline Velez

Researchers have debated including a middle response in questionnaire scales, and what selection of that response option represents. This study determined whether the number of neutral responses to attitudinal survey items is related to an item's readability, its clarity, and the respondent's ambivalence toward the attitude.

The results showed the number of neutral responses was significantly correlated with the Flesch-Kincaid Grade Level index, average letters per word, and average number of syllables per word. The results also showed that as the perceived clarity of an item increases, the number of neutral responses decreases. Partial support was found for the theory that respondents who are ambivalent about an attitude are more likely to select the neutral response.

These findings have important implications both to researchers, as well as organizations that conduct surveys on their employee population. The findings also contribute to the wealth of literature on the assessment of attitudes.

ACKNOWLEDGMENTS

There are several people who contributed to the completion of this study. I wish to give special thanks to the members of my committee: Dr. Howard Tokunaga, Dr. Lori Bartels, and Dr. Steven Ashworth for all of their suggestions and guidance throughout the process. Their patience and cooperation under such tight deadlines was definitely appreciated.

In addition, I would like to thank my family and friends who dedicated their time and energy to be my supporters, my peer reviewers, and my motivators. All of your comments and questions not only made this a better paper, it also kept me going when the light at the end of the tunnel seemed so far away. Now, you will no longer have to ask how the thesis is coming!

Finally, I would like to thank the organization used in this study. Their cooperation and support allowed this project to flow smoothly. Without their cooperation, research such as this would not be possible.

TABLE OF CONTENTS

LIST OF TABLES

The Neutral Response on Attitudinal Measures:

An Attribute of the Item

Pauline Velez

San Jose State University

Running Head: NEUTRAL RESPONSE

Footnotes

Requests for reprints should be sent to Pauline Velez, Department of Psychology, San Jose State University, San Jose, California 95192

Abstract

Researchers have debated including a middle response in questionnaire scales, and what selection of that response option represents. This study determined whether the number of neutral responses to attitudinal survey items is related to an item's readability, its clarity, and the respondent's ambivalence toward the attitude. The results showed the number of neutral responses was significantly correlated with the Flesch-Kincaid Grade Level index, average letters per word, and average number of syllables per word. The results also showed that as the perceived clarity of an item increases, the number of neutral responses decreases. Partial support was found for the theory that respondents who are ambivalent about an attitude are more likely to select the neutral response. These findings have important implications both to researchers, as well as organizations that conduct surveys on their employee population. The findings also contribute to the wealth of literature on the assessment of attitudes.

The Neutral Response on Attitudinal Measures: An Attribute of the Item

The concept of an attitude is one of the most distinctive and indispensable concepts in the field of psychology (Zimbardo, Ebbesen, & Maslach, 1990). Similar to other abstract creations, an attitude is a psychological construct (Henerson, Morris, & Fitz-Gibbon, 1987). Attitudes typically consist of three aspects; cognitive, affective, and behavioral. Therefore, measuring and evaluating an attitude is not a simple task.

The measurement of attitudes can be traced through relevant literature to as early as the 1920's (Thurstone & Chave, 1929). One of the most common methods used to measure attitudes has been the standard, written questionnaire. Most attitudinal questionnaires use a summated scale for measurement purposes. The Likert summated scale, developed by Rensis Likert, has been widely used since the early 1930's. In fact, the scale is so popular it would be difficult to find a social science journal in which at least one article did not report the measurement of attitudinal data using a Likert scale (Ryan, 1980). The purpose of this study is to investigate the relationship between one aspect of Likert surveys, the existence of a midpoint, and the readability of attitudinal survey items. In addition, this study examines the relationship between attitudinal ambivalence and the selection of the neutral response.

The Likert scale normally consists of an odd number of points. Likert used five-point scales in his early research. But since then, researchers have used scales with more points and scales with fewer points (Ryan, 1980). Some researchers believe

that when measuring well developed attitudes, using a scale with more points allows the respondents to select the most appropriate option. In other words, offering scales with very small gradations ensures that the desired response is available to choose. Precision, in terms of attitudinal measurement, increases as the number of response choices increases. Therefore, by offering more points on a scale, the researcher is able to examine varying levels of attitude intensity, and the scale's reliability increases.

The labels used to distinguish each point on the scale depend on the survey items. One such aspect is the wording of the item. In the item, "How satisfied are you with your medical benefits?", the labels would appropriately be labeled "very satisfied," "satisfied," "neither satisfied nor dissatisfied," "dissatisfied," and "very dissatisfied." Researchers determine during the questionnaire construction phase which labels are most appropriate for each point. For example, the item "When choosing an employer, the support management provides employees in resolving work/family conflicts is important" includes a five-point scale including "strongly agree," "agree," "neither agree nor disagree," "disagree," and "strongly disagree." Because of the structure of the item, it would be ineffective and inappropriate to use "very comfortable-very uncomfortable" or "very satisfied-very dissatisfied" labels.

The midpoint of this or any scale can represent a variety of opinions. For example, the midpoint can be labeled "undecided," "don't know," "no opinion" or "neutral" (Ryan, 1980). In the example of work/family conflicts, a neutral response

may mean "no opinion." Determining which label to use for the midpoint or whether to even use a midpoint should be based on the theoretical construct the researcher is attempting to measure.

Many arguments have been made both for and against the use of a midpoint. One of the main arguments in favor of offering a midpoint on a scale is that there are individuals who favor the middle position, and forcing them to choose a polar alternative will contribute to some form of random or systematic error (Schuman & Presser, 1981). As early as 1944, Rugg and Cantril argued that offering a midpoint allowed for additional gradation of opinions (1944). The inclusion of the midpoint gives the respondent more options, in addition to the polar extremities of, for example, agree and disagree. In addition, some researchers believe that offering a midpoint may make respondents more comfortable when selecting a response option (Nunnally, 1967).

One of the main arguments against offering a midpoint is that it increases the amount of error in survey data. To begin with, the inclusion of a midpoint increases the probability of response styles (Nunnally, 1967). For example, individuals who do not want to participate in the survey may answer each item using the middle alternative. In this instance, the individual differences in the responses to the survey items may not truly reflect the attitude in question. Similar to the concept of central tendency (Cascio, 1991), the data is reflecting a response style of the individual.

Lastly, many researchers who are against offering a midpoint believe that people do lean toward one direction, and that the midpoint attracts those individuals who do not feel strongly about the issue at hand (Schuman & Presser, 1981).

If a researcher decides to offer a midpoint, an accurate understanding of the underlying concepts used by respondents who select the midpoint on attitudinal surveys is necessary. A great number of studies were conducted prior to 1950 on the middle alternative response. Research on the subject declined considerably from 1950 to the mid 1980's. In the last sixteen years, a number of studies have begun to examine the middle alternative response and respondents who select the middle alternative (Bishop, 1987; Converse, 1976; Coombs & Coombs, 1976; Klopfer & Madden, 1980; Presser & Schuman, 1980; Schuman & Presser, 1981).

Research has been conducted to test whether certain item or respondent characteristics increases the number of individuals who select a certain response option (Coombs & Coombs, 1976; Klopfer & Madden, 1980; Schuman & Presser, 1980). Klopfer and Madden (1980) found that individuals who are ambivalent about an attitude are more likely to select the midpoint. Bogart (1967) illustrated that items which are sensitive in nature, multidimensional, or unclear may affect the response option that is chosen. For example, if a respondent believes that an attitudinal item addressing the subject of ethnicity is a sensitive issue, the respondent may select the midpoint that is labeled "neither agree nor disagree". By selecting this point, the

respondent may feel safe. In addition, Bogart found that items that are unclear or confusing may result in more respondents selecting the midpoint.

Item Readability

Terris (1949) was one of the first researchers to examine the level of language difficulty for a set of poll items. He concluded that some items were too complex for certain sectors of the population. At approximately the same time, Klare (1950) concluded that questions using more difficult language had a higher frequency of "no opinion" responses.

In the 1950's, Rudolf Flesch introduced an objective formula to assess the level of language difficulty for written pieces. According to Flesch (1948), the more words in each sentence and the more syllables in each word, the less readable a passage. Immediately, Flesch's "ease of reading" formula became the best known measure of language complexity or comprehension difficulty. The formula has been used to measure the readability level of various items including newspaper reports, government publications, materials for adult education and children's books (Flesch, 1948). Researchers have also applied the formula to survey items (Converse, 1976).

The practicality of using readability statistics to assess the level of language complexity on survey items seems apparent. Razek and Cone (1981) found that material requiring more effort on the part of the reader to understand resulted in greater frustration and anger, which resulted in the reader feeling negative about the

written piece. In order to eliminate this confound, one should make any written piece as simple and concise as possible so that the intended audience understands.

To assess the language complexity of a piece of written work, one must decide which readability statistics to use. For example, Flesch's Reading Ease formula examines the average sentence length and the number of syllables per 100 words, generating a score between 0 and 100. The lower the score, the more difficult the writing is to read. Flesch provides a table in order to interpret scores. For example, a written piece that receives a score in the range of 60-70 requires a reading level comparable to a 7th or 8th grader. On the other hand, a piece that receives a score in the 0 to 30 range has a very difficult reading level. In order to understand a piece within that range, the individual would need a college level or higher reading grade level.

In addition to Flesch's Reading Ease formula, there is also the Flesch-Kincaid Grade Level index. This formula looks at the average number of words per sentence and the average number of syllables per word. The score is reported as an approximate reading grade level. A readability score of between the 8th and the 10th grade is considered the most effective because it would be understood by most Americans since most Americans possess between an 8th and 10th grade reading level ability.

When deciding between usage of the Flesch Reading Ease formula and the

Flesch-Kincaid Grade Level index, the main factor the researcher must consider is the length of the writing that will be assessed. The Flesch Reading Ease formula is appropriate for longer pieces of writing since it examines the number of syllables per 100 words. For shorter items, such as attitudinal survey items, the Flesch-Kincaid Grade Level index is more appropriate since the formula examines the number of syllables per word regardless of the number of words.

Other readability statistics include the average number of words per sentence, the average letters per word, and the average number of syllables per words. All of these factors affect the readability of a written piece. If readability statistics show that a written piece is hard to understand, then the probability of the reader being confused increases. This same rationale applies to attitudinal survey items. When attitudinal items are confusing to a respondent, he or she may be more likely to endorse the neutral response.

Coombs and Coombs (1976) examined whether item ambiguity or respondent uncertainty resulted in a higher amount of "no opinion" responses. The researchers concluded that certain characteristics of questions, such as question wording or content, increase the number of "no opinion" responses because they pose more cognitive difficulty. Therefore, if item ambiguity exists, it could prove to be a confound in attitudinal measures.

What response option will a respondent select when an item is ambiguous and

an attitudinal measure does not include a "no opinion" option? If the attitudinal survey consists of a five-point Likert scale, with a middle alternative response, then it seems reasonable that respondents who are confused by an item will select the most neutral response. In this instance, the use of the middle alternative response signifies a lack of opinion due to respondent confusion.

Thus far, no research has addressed the relationship between item readability and the selection of the neutral response. This study addresses this gap in the literature.

The Present Study

The current study examines attitudinal items on an organization's annual employee opinion survey. This survey measures employee attitudes on a variety of workplace issues including pay, benefits, job satisfaction, career opportunities, company image, etc. The specific organization used in this study categorizes all items used in the employee opinion survey into one of three categories. The first category includes *standard items*, items used for many years by many nationwide organizations in the United States in their employee opinion surveys. Since many organizations use these items, benchmarking with other organizations occurs frequently. The second category consists of *core items* developed by trained experts who are employees within this specific organization. The third category includes *supplemental items* specific to each business unit within the specific organization in this study, and reviewed by, but

not necessarily developed by, the trained experts of this organization.

In addition to the above categories, the organization used in this study includes an "importance of work" item in its annual employee opinion survey. Employees are asked to select five of fourteen factors that are most important to them. The factors include benefits package, career opportunities, communications, company image, cooperative work environment, job security, management of the company, pay/income I receive, quality of supervision, recognition, training, type of work I do, working conditions, and workload. The factors chosen do not have to be the ones the employee is most or least satisfied with, but rather those factors that are most important in determining how the employee feels about working for the company. In fact, the organization initially included the item to examine what factors affect or "drive" overall job satisfaction. Therefore, by examining the factors chosen in relation to the responses on the individual items that correspond to the factors, the organization is aware of those topic areas that affect job satisfaction.

Most of the items on the employee opinion survey can be categorized into one of the factors. In other words, in addition to having a benefits package factor in the "importance of work" item, there would also be one specific question on the survey that addressed some specific aspect of the benefits offered, such as "How satisfied are you with your medical benefits package"? In fact, each factor has one to four items that address the attitude.

The purpose of the present study is to examine whether the readability level of attitudinal survey items affects the number of neutral responses associated with the item. The hypotheses to be investigated are as follows:

1. As the reading grade level of an item increases, the number of neutral responses associated with the item increases.
2. As suggested by Coombs & Coombs (1976), as the perceived clarity of an item, based on expert rater ratings, decreases, the number of neutral responses associated with the item increases.
3. According to Klopfer and Madden (1980), individuals who are ambivalent about an attitude are more likely to select the middlemost response. Therefore, for this specific organization, when an employee endorses a work factor, the endorsement signifies that the factor is one of the most important determinants of how the employee feels about working for the company. Therefore, the employee is more likely to have a definite attitude about that factor, and will be less likely to select the neutral response on the corresponding item within that work factor.

If a relationship between the number of neutral responses and an item's reading grade level does exist (hypothesis 1), it would suggest that researchers contribute to the number of neutral responses by the kinds of questions they ask. Further, in order to minimize this type of error, researchers can assess the language difficulty of the

items as part of the questionnaire construction phase. Items with a higher reading grade level may need to be reworded to ensure the majority of respondents will understand the item.

If a relationship is found between an item's clarity and the number of neutral responses associated with that item (hypothesis 2), it would further substantiate the belief that respondents are more likely to select the neutral response when the item is unclear or not understood. In addition, it would further support any analyses that have found a direct relationship between reading grade level and the number of neutral responses.

Lastly, if employees who endorse a work factor are less likely to use the neutral response (hypothesis 3), it would further substantiate Klopfer and Madden's (1980) claim that individuals who are ambivalent about an attitude are more likely to the select the midpoint. Therefore, implying that organizations need to measure the evaluative aspect of an employee's attitude, whether favorable or unfavorable, as well as the importance of the attitude to the employee.

METHOD

Survey Respondents

Data for this study was obtained from employees of a financial service organization in the United States. The respondents represented one business unit within the organization. The data from each year represents approximately 17,000 to

45,000 employees, including management (see Table 1).

The data was taken from databases containing the results of the organization's annual employee opinion survey from the years 1988 through 1991. Respondents were not given compensation for their involvement. All respondents were treated in accordance with the ethical standards of the American Psychological Association.

Materials

Items from the organization's annual employee opinion survey, from the years 1988 through 1992, were utilized. Since many of the items were used throughout the four years, this study only examined each item the first year it was used. One hundred and fifteen items were used.

The topic areas of the employee opinion survey include customer focused quality, training, performance management, communication, empowerment, work involvement, teamwork/cooperation, satisfaction with the job, job security, working conditions, workload, immediate manager or supervisor, workforce diversity, pay, benefits, recognition, company image, and leadership. There are typically one to four specific items that address each topic area on the survey.

Procedures

All items used between the years 1988 to 1992 were analyzed in terms of readability statistics using the reference software, Grammatik IV. The readability statistics that were analyzed included the Flesch-Kincaid Grade Level index, the

Table 1

Number of Employees Surveyed Each Year

Year Surveyed	Number of Employees Surveyed
1988	40663
1989	42976
1990	43118
1991	17155

average number of words per sentence, the average letters per word, and the average number of syllables per word.

All 115 items were rated on a ten-point scale in terms of its clarity by a team of expert raters. Clarity was defined as how well an item is written. The rater assessed whether the content of an item was clear or ambiguous. In order to assess the reliability of the ratings, intraclass correlations (Shrout and Fleiss, 1979) were performed.

Data from the computerized databases containing the results of the employee opinion surveys conducted between 1988 and 1991 were accessed to determine the number of neutral responses for each item. The percentage of neutral responses were used in all analyses. For those items that appear more than once in the four year span, the percentage of neutral response was based on the first year the item appeared on an employee opinion survey.

Lastly, in order to examine the relationship between attitudinal ambivalence and the selection of the neutral response, data from the 1990 employee opinion survey was utilized. Specifically, the analyses examined the endorsement or non-endorsement of the factors in the "importance of work" item in relation to the selection of the neutral response to the corresponding item.

The initial selection of the corresponding item was based on the number of items available to choose from. If a factor had only one corresponding item, then that

item was chosen. In addition, although some factors had more than one corresponding item, not all were available for this study since not all items included a truly neutral response option. For example, some items had a middle alternative responses labeled "about right." In this instance, selection of this response option actually represents a favorable opinion. Lastly, if a factor had more than one item with truly neutral responses, then a subjective judgement was made to determine which item was the most representative of the factor. Therefore, the author did not have much control over selecting the most representative item.

RESULTS

The purpose of this study was to determine if aspects of attitudinal survey items affect the number of neutral responses associated with the item. This section explores the relationships between an item's readability and an item's perceived clarity to the usage of the neutral response. In addition, this section examines how attitudinal ambivalence relates to the usage of the neutral response.

The means and standard deviations for various readability statistics and the percent of neutral responses for all 115 items are listed in Table 2.

To test the hypothesis that as the reading grade level of an item increases, as measured by the readability statistics, the number of neutral responses increases, correlations were performed between the percent of neutral responses and the Flesch-Kincaid Grade Level Index, the average number of syllables per word, the average

Table 2

Means and Standard Deviations of Readability Statistics for Survey Items (N=115)

	Mean	S.D.
Percent of Neutral Responses	19%	.08%
Flesch-Kincaid Grade Level Index	11.52	4.01
Average Number of Words Per Sentence	15.24	5.29
Average Number Letters Per Word	5.24	.96
Average Number of Syllables Per Word	1.77	.31

number of words per sentence, and the average letters per word.

The correlation matrix is given in Table 3. Specifically, the percent of neutral responses was significantly correlated with the Flesch-Kincaid Grade Level Index ($r(114)=.25$, $p < .01$), indicating that as the reading grade level of an item increases, the number of neutral responses associated with the item increases. In addition, the percent of neutral responses is significantly correlated with the average letters per word ($r(114) = .20$, $p < .05$) and the number of syllables per word ($r(114) = .22$, $p < .05$). A significant correlation was not found between the percent of neutral responses and the average number of words per sentence ($r(114) = .13$, $p > .05$).

An interesting point to note is in regards to the somewhat high standard deviation for the various readability statistics. Upon further examination, it appears that the very structure of this organization's survey contributes to the high variability in the readability statistics for the items. This particular organization structures the items in their questionnaire in terms of *standard, core,* and *supplemental items.*

As a follow-up analysis, the 1992 survey, which consisted of 60 *standard, core,* and *supplemental items,* was examined. Follow-up analyses showed that the categories of items were significantly different in terms of reading grade level, average number of words per sentence, average letters per word, and average number of syllables per word (see Table 4). After finding differences between the three categories of items in terms of reading grade level [$F(2,57) = 10.99$, $p < .001$, $N = 20$], Tukey tests, found

Table 3

Correlation Matrix for Percent of Neutral Response and Various Readability Statistics

	Percent of Neutral Responses	Flesch - Kincaid Grade Level	Avg. # of Words per Sentence	Avg. # of Letters per Word
Flesch-Kincaid Grade Level	.25**			
Avg # of Words per Sentence	.13	.42***		
Avg # of Letters per Word	.20*	.80***	.07	
Avg # of Syllables per Word	.22**	.89***	.00	.86***

Note: * = $p < .05$ ** = $p < .01$ *** = $p < .001$

that *standard* items had a lower mean reading grade level than *core* (N = 20) or *supplementai* items (N = 20). Significance was also found in terms of the number of words per sentence [F (2,57) = 14.60, p < .001] (see Table 4), with Tukey tests finding that *standard* and *core* items had a lower average number of words per sentence than *supplemental* items.

The third analysis of variance examined mean differences in the average letters per word for each category of items. Significant differences were found [F (2,57) =7.03, p < .01] (see Table 4). By examining mean differences between categories using the Tukey test, it was found that *standard* items, did in fact, have a lower average of letters per word, compared to *core* and *supplemental* items.

The last analysis of variance examined whether there were any mean differences between the categories of items in terms of the average number of syllables per word. Overall, a significant difference was found [F (2,57) = 6.54, p < .01] (see Table 4). Using the Tukey test to analyze mean differences, it was found that *standard* items had a lower average number of syllables per word than either *core* or *supplemental* items.

These follow-up analyses showed that a high amount of variability in the reading grade level of survey items can exist even within one survey. Specifically, the organization used in this study illustrates that level of readability can be a function of who develops the survey items. Looking at these categories of items, those items

Table 4

Descriptive Statistics for *Standard, Core,* and *Supplemental* Items

	Standard Items N=20		*Core* Items N=20		*Supplemental* Items N=20	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
Flesch-Kincaid Grade Level Index	7.15	3.27	10.50	3.90	12.00	2.79
Average Words per Sentence	12.55	2.54	12.60	4.02	17.85	3.94
Average Letters per Word	4.37	1.51	5.37	.85	5.23	.67
Average Number of Syllables per Word	1.49	.28	1.78	.31	1.73	.22

developed by trained experts and used by many organizations required the lowest reading grade level and were typically short and concise. On the other hand, the questions developed by untrained employees in the organization required the highest reading grade level and were long in structure. Untrained personnel tend to develop items that require a higher reading level, and are ultimately more ambiguous. This supports the recommendation that organizations, when designing "customized" surveys, should pay particular attention to the design phase of the questionnaire.

In order to investigate the hypothesis that as perceived clarity of an item decreases the number neutral responses increases, a team of seven experts rated each item on a ten-point scale in terms of its clarity, with clarity being defined as how well an item is written. The reliability of the ratings was assessed using intraclass correlations. The reliability coefficient of the expert ratings proved to be acceptable (ICC (2,7) = .75).

The clarity of an item was positively and significantly correlated with the percent of neutral responses ($r(114) = .26$, $p < .01$). Since an increase in the clarity rating indicates an item is less clear, a positive correlation symbolizes that as the clarity of an item decreases, the number of neutral responses increases.

To test the hypothesis that individuals who are ambivalent about an attitude are more likely to select the midpoint, separate chi square statistics were computed using the item that best corresponded to each factor in the "importance of work" item. Five

of the ten factors proved to be significant. The significant factors addressed the following topic areas; career opportunities [χ^2 (1) = 212.65, p < .001, phi = .07], cooperative work environment [χ^2 (1) = 11.71, p < .01, phi = .02], recognition [χ^2 (1) = 100.56, p < .001, phi = .05], type of work [χ^2 (1) = 26.59, p < .001, phi = .03], and management of the company [χ^2 (1) = 9.13, p < .01, phi = .02] (see Table 5). For each of these factors, it was found that employees who did not endorse a specific factor were more likely to select the neutral response on the corresponding item.

One factor, the workload factor, had a significant chi-square, but the phi coefficient showed it was significant in the wrong direction [χ^2 (1) = 27.43, p < .001, phi = .03] (see Table 5). In other words, employees who did not endorse the workload factor were less likely to select the neutral response.

Four chi-square analyses were not significant. Those analyses addressed the following topics; communications [χ^2 (1) = .18, p = .67, phi = .00], company image [χ^2 (1) = 3.31, p = .07, phi = .01], training [χ^2 (1) = .12, p = .73, phi = .00], and working conditions [χ^2 (1) = .09, p =.77, phi = .00] (see Table 5). Therefore, the overall conclusion is that there was partial support for the hypothesis that employees who endorse a factor will be less likely to select the neutral response.

DISCUSSION

The purpose of this study was to determine if aspects of attitudinal survey items affect the number of neutral responses associated with the item. Specifically,

Table 5

Endorsement of Factor and Selection of Neutral Response for Significant Factors

Selection of Neutral Response

	Work Factor					
	Career Opportunities	Cooperative Work Environment	Recognition	Type of Work	Manage-ment of the company	Work-load
Did not Endorse Factor	7217 (30.3%)	3856 (11.7%)	5724 (17.6%)	1479 (5.5%)	6375 (17.7%)	7720 (23.0%)
Endorsed Factor	4642 (24.0%)	1065 (10.5%)	1402 (13.4%)	698 (4.3%)	1160 (16.2%)	2403 (25.5%)
χ^2	212.65	11.71	100.56	26.59	9.13	27.43
df	1	1	1	1	1	1
sig	.000	.001	.000	.000	.003	.000
Phi Coefficient	-.07	-.02	-.05	-.03	-.02	+.03

Table 5 (continued)

Endorsement of Factor and Selection of Neutral Response for Non-Significant Factors

Selection of Neutral Response

	Work Factor			
	Communica-tions	Company Image	Training	Working Conditions
Did not Endorse Factor	8727 (23.1%)	4653 (14.1%)	5953 (16.8%)	3457 (10.9%)
Endorsed Factor	1204 (22.8%)	1370 (13.4%)	1309 (17.0%)	1231 (10.8%)
χ^2	.18	3.31	.12	.09
df	1	1	1	1
sig	n.s.	n.s.	n.s.	n.s.
Phi Coefficient	+.00	-.01	+.00	-.00

this study addressed the relationships between an item's readability and perceived clarity to the selection of the neutral response. In addition, this study addresses the relationship between attitudinal ambivalence and selection of the neutral response. This section will discuss the findings in more detail, as well as discuss the general implications the findings have on questionnaire construction and attitudinal assessment. Finally, the author discusses the limitations of this study, as well as ideas for future research.

Is Readability Related to the Number of Neutral Responses?

The first hypothesis, that as the reading grade level of an item increases, the number of neutral responses associated with the item increases, was supported. Specifically, the higher the reading grade level of an item, the more neutral responses it will have. In addition, the more letters per word and the more syllables per word, the more neutral responses.

Since items with a high reading grade level increase the probability of the respondent not understanding the items, the respondent is more likely to select the neutral response, thereby creating a response tendency which adds a certain amount of error to the study. In such a case, the researcher can no longer assume that the data in the middle alternative response option truly represents the neutral position of the respondents. Instead, some of the data may represent respondent confusion based on lack of understanding of the item.

Does Perceived Clarity Effect the Number of Neutral Responses?

Another gauge used to assess readability is perceived clarity. The second hypothesis, that as perceived clarity of an item decreases, the number of neutral responses associated with the item increases, was supported. Perceived clarity was defined as how well an item is written, and was based on ratings assigned by experts. The finding indicates that as the clarity of an item decreases, the number of neutral responses increases. This further substantiates the belief that respondents are more attracted to the middle alternative response when the item is unclear or not understood.

Attitude Ambivalence and the Neutral Response

The third hypothesis, that employees who endorse a work factor will be less likely to select the neutral response on the corresponding item, was partially supported. Five of the ten chi-square analyses proved significant. The analyses that were significant included the following topics; career opportunities, cooperative work environment, recognition, type of work, and management of the company. The workload factor was significant, but in the wrong direction. On the other hand, the analyses for factors that addressed company image, training, communications and working conditions were not statistically significant.

From a conceptual point of view, the fact that four of the factors were not significant could be indicating that there are varying levels of intensity between attitudes. In other words, employees could feel very strongly about career

opportunities, cooperative work environment, recognition, type of work, and management of the company. On the other hand, issues such as communications, company image, training and working conditions may be important to employees but to not the same degree as the former. Therefore, the levels of intensity enter into the overall picture, and become a variables that needs to be considered along with attitudinal ambivalence.

Although partial significance was found, the measures used may not have been the most effective or valid. As mentioned earlier, the author was limited by the number of items to choose from that best correspond to the factor because of the design of the questionnaire, and, in some instances, deciding which item to include was more of a subjective decision.

In general, it is possible that selection or non-selection of the neutral response on the corresponding item is not the best measure of ambivalence. It could be that the one item does not completely address why a particular employee endorsed the subsequent work factor. For example, a respondent may feel working conditions is one of the most important factors, but the corresponding item may address an issue that was not relevant to the respondent, such as access to computer equipment. Instead, the respondent may have endorsed the factor because the amount of noise or lighting is a significant aspect. Therefore, it is possible that the corresponding items examined were not the most valid measures for all respondents.

Another related issues is that the corresponding item may have been too general. In which case, the item was too ambiguous for the respondent, and the respondent selected the middle alternative response due to ambiguity, even though the factor was one of the most important factors to the respondent.

Specifically, although the workload was significant, it was in the wrong direction. One possible explanation could be that although the respondent feels workload is an important factor, the respondent may feel fearful of expressing either satisfaction or dissatisfaction with his or her current workload out of fear that his or her workload will change once management receives this information. Therefore, selection of the middle alternative response provides a "safe" alternative.

In terms of company image, the item used in the analysis assesses the level of satisfaction the respondent has with the overall service the company provides its customers. Again, one possible explanation is that although company image is an important factor for many respondents, the respondents may not feel that the level of service the company provides its customers is a direct measure of company image. If this were the case, this would not be a representative item. Instead, the respondent may feel that whether or not the company is involved in community-sponsored activities has a larger impact on the company image.

Implications of This Study

This study has provided additional insight to the age old debate of whether or

not to include a middle alternative response. The study does provide some significant findings. Particularly, the reading grade level of survey items does contribute to the number of neutral responses associated with items, and, ultimately, does create an amount of error. This is possibly the greatest contribution of this study considering the lack of literature that exists on the selection of the neutral response as it relates to the readability of survey items.

From the perspective of readability, this study has illustrated that the interpretation of middle alternative data must include some discussion of the readability of the survey, especially if the survey did not examine this issue during the design phase. This becomes more important to those researchers who do not analyze middle alternative data. Decisions whether or not to include the middle alternative on surveys and to analyze middle alternative data could skew the results.

The largest implication of this study suggests that surveys need to be assessed in terms of readability to ensure the questionnaire will not create a higher level of cognitive difficulty for the intended audience. It appears as if, as researchers, the temptation exists to overestimate the degree to which our audience is informed and able to understand survey items. The same holds true for organizational surveys. This study highlights the importance of considering readability when piloting and proofing questions, even when trained experts develop the questions. Although the specific organization used in this study does employ experts who review the questions on the

employee opinion survey, some of the error could have been eliminated had the items been more readable. As more and more organizations move to surveying employees and customers, the implications of this study are magnified. Organizations need to consider readability in the piloting process of their questionnaires.

Limitations

Most of the limitations in this study relate to problems of working with archival data. For this particular study, it was not possible to ask follow-up survey items to address attitude intensity, but future research in this area could benefit from this. In other words, in addition to asking the standard survey item, there could be one additional follow up item that addresses the level of attitude intensity.

Future Research

Although research does exist that examined the relationship of the level of education, sex, age, and race and the selection of the middle alternative response, thus far, the results have been inconclusive. If further insight is desired, the possibility of replicating such studies still exists, but more substantive findings might emerge by examining segments of respondents. Segmentation analyses, popular in the market research field, could provide the interaction between education, sex, age, race and the selection of the neutral response, and might provide insight into how much variation in neutral responding is due to individual differences.

Conclusions

This study contributes to the literature both in terms of the neutral response, as well as the readability of attitudinal survey items. Specifically, this study shows that items with a high reading grade level do, in fact, have more neutral responses. The implications are generalizable to individuals who conduct any type of survey, whether or not the response scale includes the neutral response, because the readability of a survey can ultimately affect the data, and any interpretation of the data. Readability, from this stand point, becomes an important issue in questionnaire design.

In addition, this survey raises important considerations and thought-provoking questions about attitudinal measurement from a more conceptual point of view. Particularly, this study examines the methods in which an organization measures attitudes, as well as reinforces the fact that evaluating and measuring attitudes is not a simple task, especially considering the intrinsic complexity of attitudes and the various aspects that encompass attitudes.

REFERENCES

Bishop, G. F. (1987). Experiments with the middle response alternative in survey questions. Public Opinion Quarterly, 51, 220-232.

Bogart, L. (1967). No opinion, don't know, and maybe no answer. Public Opinion Quarterly, 31, 332-345.

Cascio, W. F. (1991). *Applied Psychology in Personnel Management.* New Jersey: Prentice Hall.

Converse, J. M. (1976). Predicting no opinion in the polls. Public Opinion Quarterly, 40, 515-530.

Coombs, C. H. & Coombs, L. C. (1976). "Don't know": Item ambiguity or respondent uncertainty? Public Opinion Quarterly, 40, 497-514.

Flesch, R. (1948). A new readability yardstick. Journal of Applied Psychology, 32, 221-233.

Henerson, M. E., Morris, L. L., & Fitz-Gibbon, C. T. (1987). How to measure attitudes. Newbury Park, California: Sage.

Klare, G. R. (1950). Understandability and indefinite answers to public opinion questions. International Journal of Opinion and Attitude Research, 4, 91-96.

Klopfer, F. J. & Madden, T. M. (1980). The middlemost choice on attitude items: Ambivalence,neutrality, or uncertainty? Personality and Social Psychology Bulletin, 6, 97-101.

Nunnally, J. C. (1967). *Psychometric Theory*. New York: McGraw Hill.

Presser, S. & Schuman, H. (1980). The measurement of a middle position in attitude surveys. Public Opinion Quarterly, 44, 70-85.

Razek, J. R. & Cone, R. E. (1981). Textbooks - and empirical study. Journal of Business Communication, 18, 33-40.

Rugg, D. & Cantril H. (1944). The wording of questions. In H. Cantril (ed.), Gauging Public Opinion (p. 33). Princeton, NJ: Princeton University Press.

Ryan, M. (1980). The Likert scale's midpoint in communications research. Journalism Quarterly, 305-313.

Schuman, H. & Presser, S. (1981). Questions and answers in attitude surveys. San Diego, California: Academic Press Inc.

Shrout, P. E. & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. Psychological Bulletin, 86, 420-428.

Terris, F. (1949). Are poll questions too difficult? Public Opinion Quarterly, 13, 314-319.

Thurstone, L. L., & Chave, E. J. (1929). The measurement of attitude. Chicago: Univ. Chicago Press.

Zimbardo, P. G., Ebbesen, E. B., & Maslach, C. (1990). Influencing attitudes and changing behavior. New York: McGraw-Hill.

SAN JOSE STATE UNIVERSITY

Appendix A

Office of the Academic Vice President • *Associate Academic Vice President* • *Graduate Studies and Research*
One Washington Square • San José, California 95192-0025 • 408/924-2480

To: Pauline Velez
495 Sioux Lane
San Jose, CA 95123

From: Serena W. Stanford Serena W. Stanford
AAVP, Graduate Studies and Research

Date: December 7, 1992

The Human Subjects-Institutional Review Board has reviewed and approved your request for exemption from Human Subjects Review for the proposed study entitled:

> "The Neutral Response on Attitudinal Measures: An Attribute of the Item"

Provided that there are no changes in the procedure proposed, you may proceed with this study without further review by the Human Subjects-Institutional Review Board. You must notify the Human Subjects-Institutional Review Board of any changes in the subject population or procedure for this study

I do caution you, however, that Federal and State statutes and University policy require investigators conducting research under exempt categories to be knowledgeable of and comply with Federal and State regulations for the protection of human subjects in research. This includes providing necessary information to enable people to make an informed decision regarding participation in your study. Further, whenever people participate in your research as human subjects, they should be appropriately protected from risk. This includes the protection of the confidentiality of all data that may be collected from the subjects. If at any time a subject becomes injured or complains of injury, you must notify Dr. Serena Stanford immediately. Injury includes but is not limited to bodily harm, psychological trauma and release of potentially damaging personal information.

Please also be advised when people participate in your research as human subjects, each subject needs to be fully informed and aware that their participation in your research project is voluntary, and that he or she may withdraw from the project at any time. Further, a subject's participation, refusal to participate or withdrawal will not affect any services the subject is receiving or will receive at the institution in which the research is being conducted.

If you have questions, please contact me at 408-924-2480.